Diffuse reflectance spectroscopy accurately discriminates early and advanced grades of fatty liver in mice

Eduardo J. Arista Romeu
Galileo Escobedo
Adriana Campos-Espinosa
Ivette Irais Romero-Bello
Javier Moreno-González
Diego A. Fabila-Bustos
Alma Valor Reed
Suren Stolik Isakina
José Manuel de la Rosa Vázquez
Carolina Guzmán

SPIE.

Diffuse reflectance spectroscopy accurately discriminates early and advanced grades of fatty liver in mice

Eduardo J. Arista Romeu,[a,†] Galileo Escobedo,[b,†] Adriana Campos-Espinosa,[b] Ivette Irais Romero-Bello,[b] Javier Moreno-González,[b] Diego A. Fabila-Bustos,[a,c] Alma Valor Reed,[a] Suren Stolik Isakina,[a] José Manuel de la Rosa Vázquez,[a,*] and Carolina Guzmán[b,*]

[a]Instituto Politécnico Nacional, Escuela Superior de Ingeniería Mecánica y Eléctrica, Unidad Zacatenco, Laboratorio de Biofotónica, Ciudad de México, Mexico
[b]Hospital General de México "Dr. Eduardo Liceaga," Universidad Nacional Autónoma de México, Facultad de Medicina, Unidad de Medicina Experimental, Laboratorio de Hígado, Páncreas y Motilidad, Ciudad de México, Mexico
[c]Unidad Profesional Interdisciplinaria de Ingeniería Campus Hidalgo, Instituto Politécnico Nacional, Ciudad del Conocimiento y la Cultura, San Agustin Tlaxiaca, Hidalgo, Mexico

Abstract. Nonalcoholic fatty liver disease (NAFLD) ranges from steatosis to nonalcoholic steatohepatitis and cirrhosis. Liver biopsy, considered the gold standard to diagnose NAFLD, shows significantly high rates of interobserver variability. Thus there is a need to develop tools that accurately categorize mild and advanced grades of steatosis in order to identify patients at higher risk of developing chronic liver disease. Diffuse reflectance spectroscopy (DRS) has proved to be useful in grading liver fibrosis and cirrhosis, without having been implemented for steatosis. We aim to categorize early and advanced stages of liver steatosis in a methionine–choline deficient (MCD) mouse model. C57bl/6 mice are fed either methionine–choline control or MCD diet during 2 or 8 weeks to induce mild and advanced steatosis. Liver samples are obtained and steatosis is evaluated by oil red O staining. Diffuse reflectance spectra are directly measured on *ex vivo* liver specimens, in a wavelength range of 400 to 800 nm. DRS is able to discriminate between early or advanced steatosis and healthy hepatic tissue with negligible error while showing high average sensitivity and specificity (0.94 and 0.95, respectively). Our results suggest that liver steatosis can be accurately evaluated by DRS, highlighting the importance of applied spectroscopic methods in assessing NAFLD. © *2018 Society of Photo-Optical Instrumentation Engineers (SPIE)* [DOI: 10.1117/1.JBO.23.11.115005]

Keywords: diffuse reflectance spectroscopy; steatosis; nonalcoholic fatty liver disease.

Paper 180370R received Jun. 20, 2018; accepted for publication Oct. 18, 2018; published online Nov. 29, 2018.

1 Introduction

Nonalcoholic fatty liver disease (NAFLD) includes a wide spectrum of histopathological changes characterized by abnormally high accumulation of lipids in the liver of patients with negligible alcohol intake or seronegative hepatitis C virus infection.[1] NAFLD might potentially progress through the lipid accumulation typically found in simple steatosis to nonalcoholic steatohepatitis (NASH), which in turn might lead to the development of fibrosis, cirrhosis, hepatocellular carcinoma (HCC), and liver failure.[2] It is then of great importance to perform an accurate classification of simple steatosis with the aim to identify patients at higher risk of developing NASH, cirrhosis, and HCC.[3]

Liver biopsy remains the gold standard procedure for scoring the grade of NAFLD.[4,5] However, numerous studies have consistently reported a significant intra- and interobserver variability at the time of scoring the level of intrahepatic lipid accumulation, most of all when attempting to measure mild-to-moderate grades of NAFLD.[3–8] Unfortunately, such a marked variability may result not only in underestimation of the steatosis degree but also in the incorrect identification of the disease,

which in turn might lead to lack of an accurate medical treatment for patients.[5,7,8] Therefore, it is imperative to implement tools that are complementary to liver biopsy with the aim to score different degrees of liver steatosis in an accurate manner, thus reducing the intra- and interobserver variability and increasing the chance to provide opportune diagnosis and medical treatment for patients with risk of developing chronic liver disease.

Diffuse reflectance spectroscopy (DRS) is a methodology that allows identifying the absorption and scattering patterns of biological tissue.[9,10] The absorption spectrum depends on the assortment of molecules present in the tissue, whereas the scattering pattern is mainly related to the concentration and size of those molecules.[11–13] The absorption and scattering patterns directly result from the diversity of molecules that compose a tissue, giving to each biological sample a typical optical signature also referred to as diffuse reflectance spectrum.[14]

DRS has been recently used to study lipid deposition in the liver.[15,16] In this sense, Evers et al.[15] demonstrated that human liver specimens exhibiting micro- and macrovesicular steatosis show a diffuse reflectance spectrum significantly different to that found in control liver samples without lipid depots, when studied *in vivo* and *ex vivo* at 1200 nm. A similar result has been found in murine models of NAFLD, due to lipids showing an absorption peak around 1200 nm.[17] However, there is still not

enough experimental evidence regarding the characterization of hepatic diffuse reflectance (DR) spectra in the early stage of simple steatosis, especially during mild and moderate stages of fatty liver disease.

The main objective of this work was to examine whether the liver tissue has a specific diffuse reflectance spectrum depending on the level of steatosis while also evaluating the feasibility of scoring liver steatosis by DRS as a complementary tool to liver biopsy in a murine model of NAFLD.

2 Materials and Methods

2.1 Steatosis Induction

Male C57BL/6 mice of 16 weeks of age and weighting 25 ± 5 g were obtained from the Animal Care Facilities at the Experimental Medicine Unit from the Hospital General de México and maintained under controlled conditions. Both food and water were allowed *ad libitum*. All procedures were approved by the Institutional Ethics Review Board at the Hospital General de México.

In order to develop different degrees of steatosis, mice were randomly assigned to be fed either a methionine–choline deficient (MCD) or a methionine–choline control (MCC) diet (MP Biomedicals, California, USA) during 2 (MCD2w) or 8 weeks (MCD8w). To avoid confounding factors due to the complexity of the degree of NAFLD, only animals developing simple steatosis were included in this study.

2.2 Sample Collection

After appropriate time of treatment, 2 weeks for MCD2w and 8 weeks for MCD8w, animals were anesthetized with xylazine–ketamine. Liver samples from the right lobe were obtained and preserved in sterile cold PBS for immediate DRS assessment as described below. The left lobe was divided to either be embedded in Tissue-tek OCT (Sakura Finetek, USA) and stored at $-20°C$ until assayed; or be fixed in 3.7% formaldehyde-PBS solution to be embedded in paraffin.

2.3 Hepatic Fat Content Assessment

Histological evaluation of the degree of fatty liver disease was performed in Hematoxylin-Eosin stained sections according to the nonalcoholic fatty liver disease activity score (NAS).[18] Liver fat contents were assessed in frozen sections stained with oil red O (Abcam, California, USA) and quantified by a morphometric analysis using Image J software (NIH, USA).

2.4 Diffuse Reflectance Instrumentation and Measurements

The diffuse reflectance instrumentation used in this study has been reported elsewhere.[9] Briefly, the light source employed was a light emitting diode (LED-P3WW120-120/41 115 by SiLED Int), which operates in the range of 400 to 800 nm. The spectrometer employed was a QE65000 (Ocean Optics Corp.), which operates in the region from 400 to 1100 nm with an entrance slit of 10 μm. The fiber-optic bundle (QR400-UV-VIS, Ocean Optics Corp.) is composed of two branches, one with a single fiber-optic to collect the diffusely reflected light and deliver it to the spectrometer, and the other with a circular array of six fibers that are employed for irradiation. All fibers have a diameter of 400 μm. The operation of both spectrometer and light source is performed by means of a programmed LabVIEW interface. A schematic of the instrumentation is shown in Fig. 1.

Ex vivo DR spectra were acquired from a total of 22 samples of liver tissue gathered from the same number of subjects: nine for MCC group, six for MCD2w group, and seven for MCD8w group. Measurements were carried out in a darkened room putting the probe tip manually establishing surface contact with the liver surface at different points without puncturing the Glisson's capsule at a spatial density of 22 points per square centimeter. This procedure was always conducted by the same operator to ensure that a nearly constant pressure was applied. Prior to spectral measurements of liver samples and after 15 min warm-up time, the dark spectrum was obtained by blocking the light input to the spectrometer. During data processing, the dark spectrum was subtracted from each measured spectrum using
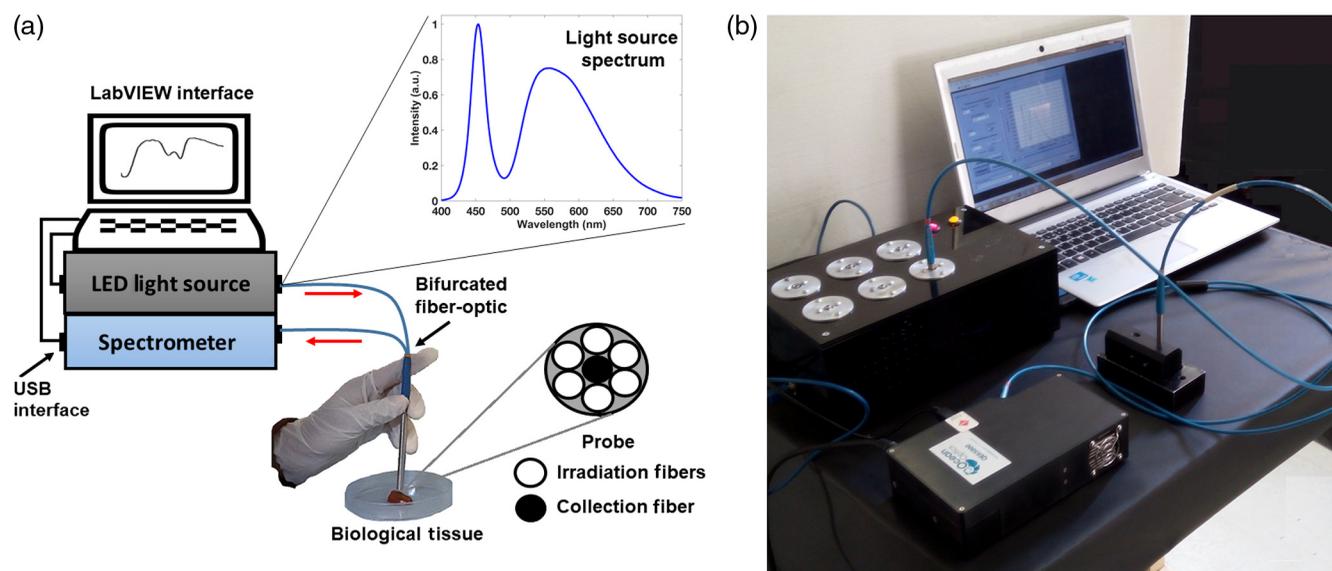

Fig. 1 (a) DRS setup: the spectrum of the excitation light source as well as the geometry of the probe is shown and (b) system showing the measurement of the diffuse reflectance standard.

appropriate software. Also, reference LED-source reflectance spectrum was measured using a reflectance standard (Spectralon WS-1, Ocean Optics Corp.) with a reflectivity above 95% for the 250- to 1500-nm range, positioning the probe in close contact with the surface of the standard and perpendicular to it. For all measurements, the radiant flux power of the light source was configured at 0.7 mW (measured at the tip of the probe). The integration time of the spectrometer was set to 100 ms. All spectra were limited to the 400- to 800-nm range. The spectrometer setting was such that its spectral resolution is equal to 0.77 nm, which yields 522 measuring points in each spectrum.

Diffuse reflectance measurements of each sample were converted to relative reflectance values using

$$R_d(\lambda) = \frac{S(\lambda) - D(\lambda)}{S_{\text{ref}}(\lambda) - D(\lambda)}, \tag{1}$$

where $R_d(\lambda)$ is the relative reflectance intensity at wavelength λ, $S(\lambda)$ is the sample intensity at wavelength λ, $D(\lambda)$ is the dark intensity at wavelength λ, and $S_{\text{ref}}(\lambda)$ is the reference (source) intensity at wavelength λ.

2.5 Statistical Analysis

A total of 428 DR spectra were acquired: 153 spectra for MCC group, 143 for MCD2w group, and 132 for MCD8w group. For every group, the mean of all the measured spectra as well as its standard deviation were calculated. Areas under the curve (AUC) of all spectra in the interval from 400 to 800 nm were computed in relative units. Then, one-way ANOVA followed by Tukey *post hoc* tests were carried out to determine statistical differences between the means of the AUC of each group with a 0.05 significance level.

As a second step, a multivariate statistical analysis of principal components (PCA) was performed over the whole spectral data set. This analysis reduces the number of variables with minimum loss of information.[19] Principal components (PCs) that accounted for most of the variance (>99%) and that correspond to the eigenvalues before the cut-off point or "elbow" of the Scree plot[20] were taken for the construction of a linear discriminant analysis (LDA) classification model. A linear model was chosen due to the limited number of available observations and its better stability to variations in the statistical data.[21] A supervised classification scheme was developed using the LDA model, where the independent variables were the five most significant PCs that accounted for over 99% of the variance of the whole set of observations, and taking into account the membership to the groups—MCC, MCD2w, and MCD8w—as the classes in this classification framework. Afterward, the classification and cross-validation procedures were carried out following a leave-one-out scheme for each one of the observations, i.e., for each measured spectrum that was reduced to its five PCs, to prove the belonging of each observation to its corresponding class. To be specific, all data were used as training data and cross validation was performed excluding data "one at a time" to judge which group it should be assigned to and then verifying whether the classification was correct or not. Such framework allows the validation of the classification scheme through the determination of the confusion matrix, which was constructed to contrast the predicted versus the actual class membership; the calculation of the classification rate errors as the value in percent of misclassifications relative to the number of observations.

Also, a canonical score plot of LDA was performed to define the axes that provide and visualize maximum separation among the groups. Given the fact that the canonical variables are uncorrelated linear functions of the original set of variables, in our case, there were three groups or classes; therefore, there will be two canonical variables. Probability estimates obtained from the above classification model were then used as scores to obtain receiver operating characteristic (ROC) curves fixing in every case one state as positive (i.e., belonging to its class) and the other two as negative. The ROC curve plots the proportion of true positives (sensitivity) against the proportion of false positives (1 − specificity) for every possible value of the threshold (in our case, the posterior probability obtained from the classification model).[22] The AUC was obtained for every ROC curve and they were used for comparison of the discriminative power of different tests. Optimal values of sensitivity and specificity for each ROC curve were found using the criterion of the Youden index, which is calculated as the cut-point where the maximum of (sensitivity + specificity) − 1 occurs.[23] In this way, the effectiveness of using the classification results as predictors was evaluated, the Youden index is optimal in the sense that it is the cut-point that optimizes the predictors differentiating ability when equal weight is given to sensitivity and specificity. The whole statistical analysis was conducted using MATLAB R2016a (Mathworks, Natick, Massachusetts, USA).

3 Results

3.1 Hepatic Steatosis

Mice receiving MCC diet for 2 or 8 weeks did not exhibit any hepatic architectural alteration as it can be observed in the first row of Fig. 2(a), where typical images of Hematoxylin-Eosin stained liver sections are displayed on the left side; on the right side, characteristic images of liver sections stained with oil red are shown. Thus liver samples from mice fed with MCC diet for 2 or 8 weeks were considered as one single control group (MCC). For the MCD groups, simple steatosis was confirmed by NAS (values 1 or 2). According to the time of treatment, significantly increased liver fat was observed in mice fed MCD diet [see the second and third rows of Fig. 2(a)]. The obtained percent values of liver fat content are displayed in a bar plot in Fig. 2(b), where the error bars stand for the positive half of the standard deviation. An ANOVA statistical test was performed with the null hypothesis that the means of all groups are equal at the 0.05 significance level. The null hypothesis could be rejected by ANOVA test. However, it was still necessary to determine which one of the fat-content mean-values is different from the rest. For that reason, a Tukey multiple comparison test was applied. It was found that it is not possible to differentiate the mean fat content between the MCD2w and MCD8w groups [indicated by a line over the bars in Fig. 2(b)] while the mean fat content of the MCC group differentiates, with 0.05 significance level, from both simple steatosis groups: MCD2w and MCD8w. However, the mean size of fat droplets was significantly different among MCD2w and MCD8w (MCD2w = 17.3 ± 8.5; MCD8w 33.0 ± 7.9 μm, $p < 0.05$).

3.2 Diffuse Reflectance Spectra

Average diffuse reflectance measurements were obtained for every group. The calculated mean ± dispersion bars spectra for each group are plotted in Fig. 3(a). Although the *dispersion*
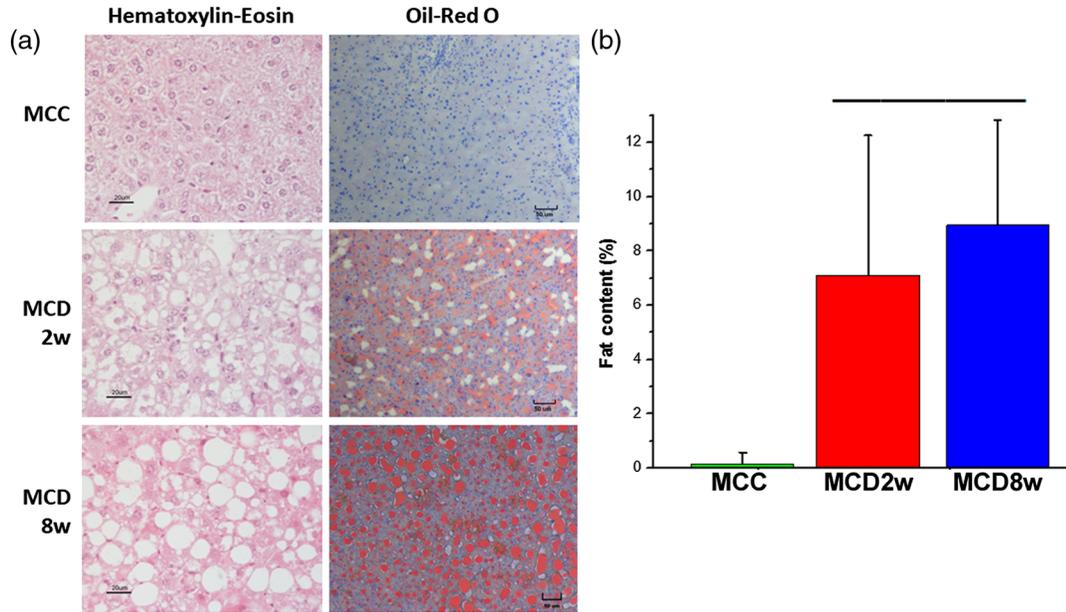
Fig. 2 (a) Images of tinctured liver samples and (b) liver fat content expressed in percent values. Mean value ± standard deviation. $p < 0.05$, ANOVA. Line over bars indicates groups, which were not significantly different (Tukey multiple comparison test).
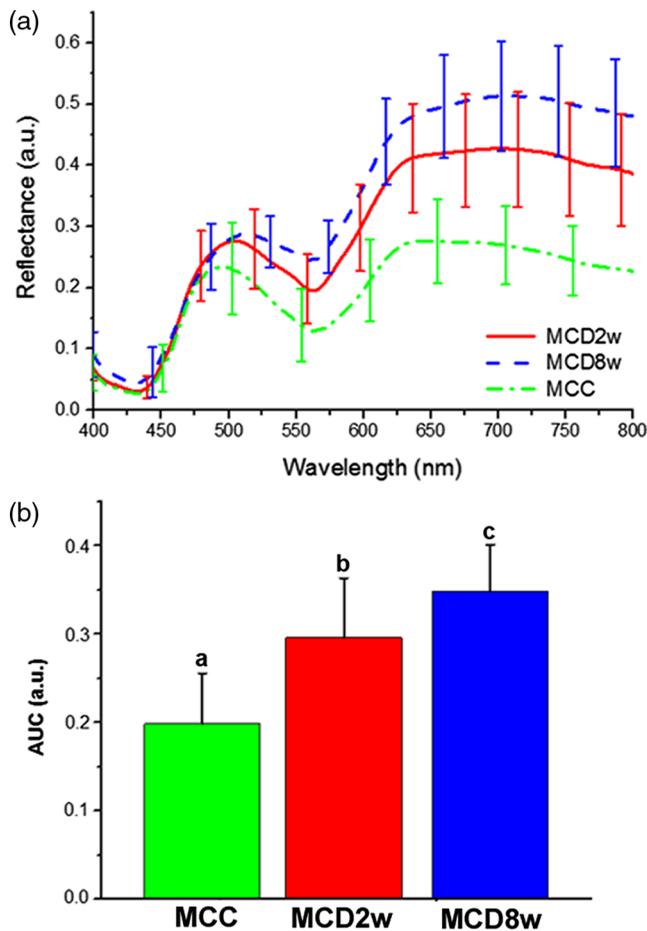

Fig. 3 (a) Mean DRS ± standard deviation from MCC, MCD2w, and MCD8w groups and (b) bar plot of the AUC ± standard deviation of the DRS from 400 to 800 nm shown in (a). Bars not sharing a letter are significantly different (Tukey test, $p < 0.05$).

bars for MCD groups partially overlap, there is evidence that reflectance spectra from different groups have intensity differences in the range from 500 to 800 nm, which become more significant as the fat content increases. Interestingly, around 560 nm, a large decrease in reflectance is observed. Also, the AUC from 400 to 800 nm seems to point to a difference between groups with simple steatosis and the control (AUC: MCC = 0.20 ± 0.06, MCD2w = 0.29 ± 0.07, and MCD8w = 0.35 ± 0.05; $p < 0.05$) as seen from Fig. 3(b); these differences increase when the AUC is calculated from 630 to 800 nm (AUC for MCC = 0.26 ± 0.06, MCD2w = 0.41 ± 0.09, and MCD8w = 0.50 ± 0.09; $p < 0.05$). The ANOVA and Tukey tests yielded that at 0.05 significance level it is possible to reject the hypothesis that the population means are equal and determined that the difference between the means of all groups is significant, thus pointing to the fact that indeed the DR spectra have distinctive differences for all groups. However, a more detailed statistical analysis will follow below to delve into these observations. Visually, the percentage of liver fat content in Fig. 2(b) and the areas of DR curves in Fig. 3(b) appear to be correlated, in fact averaging those variables for each animal a Spearman correlation coefficient of 0.836 was found.

3.3 PCA and LDA

The first five PC values, accounting for 99.97% of the variance, were estimated and selected for a supervised discriminant analysis classification. Even though equality test of group covariance matrices failed at 0.05 significance level (natural log of the determinants of each group's covariance matrix being −19.94, −18.35, and −15.72, respectively), the linear variant of the discriminant analysis was enforced because the number of observations was near the recommended lower valid number of observations for its quadratic analog, which may cause an incremented variance with variations on input data. The canonical
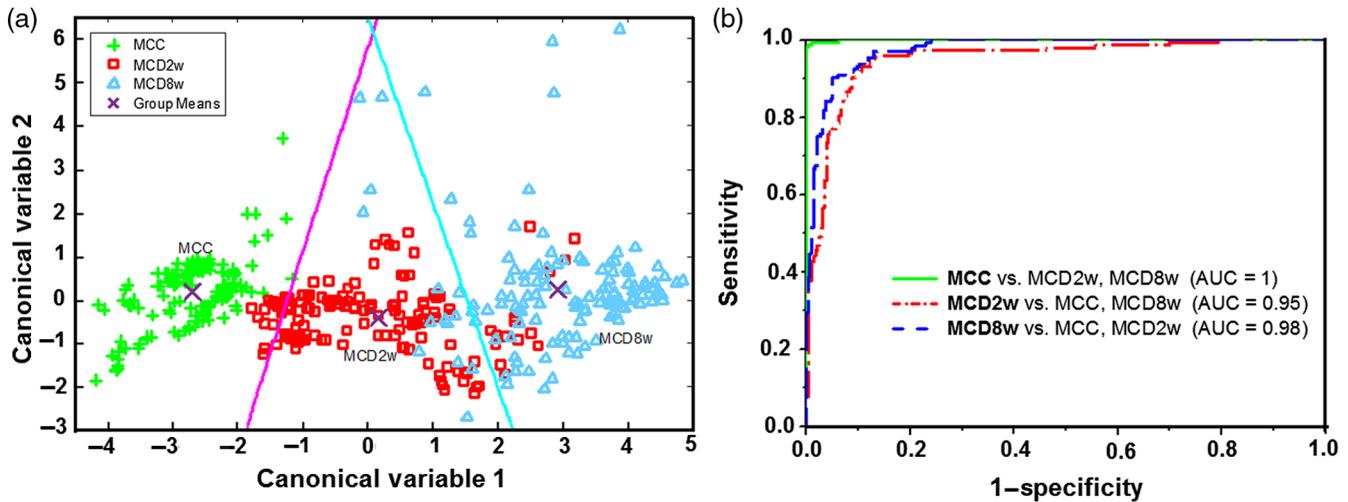
Fig. 4 (a) Canonical score plot for the LDA performed with the five first PCA values of diffuse reflectance spectra. The straight lines show the boundaries between classes obtained from the LDA model. (b) ROC curves describing the capacity of correctly assigning any of the measurements to its group. The AUC is referred for each curve.

score plot of the LDA is shown in Fig. 4(a), where the center of each group is marked with "×." The coefficients of the equations of the straight lines shown in Fig. 4(a) for the boundaries between classes were readily available from the linear discriminant classifier model. Discrimination of the healthy group from the groups suffering from fatty liver is observed. Table 1 shows the confusion matrix (actual classification of groups based on the exposure time to MCD diet versus predicted classification groups obtained from DRS measurements plus statistical analysis) for the LDA. The diagonal values (highlighted in bold font) reflect the percentage of correct group classification, where all controls were correctly assigned to control group, whereas 78.32% and 87.88% were assigned to the simple steatosis groups MCD2w and MCD8w, respectively. Values in the total line reflect the sum in percentage of all measurements assigned by the classification model to every predicted group, relative to all measurements.

Cross-validation error rates of 0.00%, 22.38%, and 12.88% were close enough to the classification error rates at 0.00%, 21.68%, and 12.12% for the respective MCC, MCD2w, and

MCD8w groups to validate the used classification model. The corresponding overall cross-validation classification error at 11.45% is calculated using the cross-validation error rate values obtained for each group corrected by the different in group membership of each group, this difference is accounted multiplying each error rate value for the prior probability for each group (number of observations in the group divided by the total number of observations) and it was in good agreement with the 10.98% classification error, which was calculated in a similar way $(21.68\% * 0.33411) + (12.12\% * 0.30841) = 10.9814\%$. Figure 4(b) shows the obtained ROC curves; the high specificity and sensitivity are observed for control (MCC) versus simple steatosis (MCD2w and MCD8w). For each ROC curve, the AUC was obtained and the optimal threshold points in the form of (sensitivity, 1 − specificity) were found according to the criterion of the Youden index as follows: (0.99, 0.00), (0.93, 0.11), and (0.90, 0.05) for MCM, MCD2w, and MCD8w groups, respectively; which gives an average value of (0.94, 0.05).

4 Discussion

The liver is mainly composed by water $(76 \pm 7)\%$, lipids $(19 \pm 11)\%$, bile $(6.6 \pm 4.5)\%$, and blood $(3.4 \pm 2)\%$ (all expressed in volume fractions). Under visible light, these components show absorption coefficients of the order of $\leq 10^{-2}$, $\leq 10^{-1}$, 10^2 to 10^0, and 10^3 to 10^1 cm^{-1}, respectively.[24] Fat and blood show similar reduced scattering coefficients of about of 10 cm^{-1}.[25,26] When the lipid fraction in the liver rises, as a consequence of steatosis progression [see Fig. 2(b)], the light scattering increases and absorption decreases in the hepatic parenchyma, producing a higher diffuse reflectance, as it can be seen in Fig. 3(a). The increasing size of the fat droplets as steatosis advances also contributes to the reflectance growth,[27] which could account for a rising of the reduced scattering coefficient of about 25% when the droplets average diameter rises from 17.3 ± 8.5 μm for the MCD2w to 33.0 ± 7.9 μm for the MCD8w.

The decrease of the spectra at 560 nm is consistent with the peak of the absorption coefficient in the healthy liver[28] as a consequence of peaks of maximal absorption for hemoglobin

Table 1 Confusion matrix for the linear discriminant model.

	Predicted group			
	MCC (number), %	MCD2w (number), %	MCD8w (number), %	Total (number), %
MCC (actual)	**(153) 100.00%**	(0) 0.00%	(0) 0.00%	(153) 100.00%
MCD2w (actual)	(16) 11.19%	**(112) 78.32%**	(15) 10.49%	(143) 100.00%
MCD8w (actual)	(0) 0.00%	(16) 12.12%	**(116) 87.88%**	(132) 100.00%
Total	(169) 39.49%	(128) 29.91%	(131) 30.61%	(428) 100.00%

Note: The diagonal values (highlighted in bold font) reflect the number of correctly classified observations and their percentages.

(Hb, Hb-O$_2$) and fat in the range of 500 to 560 nm.[24,29] Moreover, the minimal reflectance value at 560 nm in Fig. 3(a) agrees with the maximal peaks of the Hb and fat absorption coefficients.[24,29] A reduction of the absolute slope in the range of 500 to 550 nm is also noticeable as steatosis grows (-0.62 nm^{-1} for MCC, -0.30 nm^{-1} for MCD2w, and -0.04 nm^{-1} for MCD8w) that could be used as a differentiation parameter.[30] However, the uncertainties of the slope values are high, increasing with the tissue fat content for the MCD8w group. This result is consistent with the observed histological images, such as the ones depicted in Fig. 2(a), where the tissue heterogeneity increases together with the fat content. These heterogeneities should be responsible for the increasing dispersion in the slope values in the region of interest of the spectra taken at different points in the tissue.

Our present results describe the use of DRS in assessing liver steatosis in *ex vivo* samples from an experimental model of NAFLD. The approach used here was invasive, and it needs a direct contact of the probe with the sample. In this sense, it could be relevant in liver graft assessment, from both cadaveric and live donors, before liver transplantation. Nevertheless, a lesser invasive application of this tool *in vivo* is possible using endoscopic instrumentation, we have previously reported the use of DRS as an effective tool to evaluate the optical differences among normal and fibrotic liver both in an animal model of hepatic chronic disease[12] and in paraffin preserved human samples.[31] Other reports support DRS use in assessing liver pathologies.[14–17,24,32]

NAFLD is an increasingly common hepatic disease worldwide, with a spectrum that comprises from simple steatosis, a benign lipid accumulation in the hepatic parenchyma; through NASH, which besides steatosis, includes inflammation and features of oxidative stress. NASH is considered a cause to develop fibrosis, cirrhosis, HCC, and liver failure.[2] Diagnosis of simple steatosis and NASH is performed by liver biopsy; in fact, this is the only accepted tool able to discern simple steatosis from NASH. Structural and compositional changes evident in the hepatic parenchyma could be used to apply spectrometric strategies in evaluating the progression of NAFLD. Particularly, it is of great importance to assess the quality of liver grafts prior to transplantation and to perform an accurate classification of the mild and moderate grades of lipid accumulation in the liver with the aim of identifying patients at higher risk of progressing from simple steatosis to NASH and end-stage liver disease. However, these changes become increasingly heterogeneous as the disease advances. In this paper, we report data only on simple steatosis in order to avoid confusion due to the more heterogeneous compositional differences present in NASH. We excluded samples showing inflammation (only NAS $\leq$2 were included); we hypothesized that the presence of cells and molecules involved in inflammation might account for higher heterogeneity in the tissue, therefore, differences on its optical properties as well as the DRS obtained. Although we identified some subjects showing inflammation and ballooning (that were fed MCD diet during 12 weeks and were excluded from this analysis), by now we have not analyzed the differences between NASH and simple steatosis and their relationship to fat or to other factors derived from inflammation and fibrosis.

Previous work by Westerkamp et al.[17] has shown the applicability of interstitial *ex vivo* infrared DRS (around 1200 nm) in lipid quantification in murine models of fatty liver disease; obtaining for the differentiation between mild and moderate

severe/severe steatosis a sensitivity and a specificity of 86% and 81%, respectively. Here, using DRS from 400 to 800 nm, we report average sensitivities and specificities of 94% and 95%, respectively, for the discrimination between healthy, mild, and advanced steatosis. We were able to assess the presence of steatosis in the liver exposed to an MCD diet for 2 or 8 weeks. Lipid contents were associated to an increased AUC of the corresponding spectra compared to control livers [Fig. 3(b)].

The five most significant PCs provided a very good separation among hepatic steatosis and healthy livers with 99% sensibility and 99% specificity. The applicability of DRS in assessing simple steatosis was confirmed by LDA [Fig. 4(a)]. In fact, misclassification rates were low (<22%) among the groups. Application of the first five PCs allowed us to discriminate simple steatosis from healthy liver with high specificity and sensitivity as shown in the ROC curves [Fig. 4(b)]. The averaged AUC obtained for the ROC curves was 0.98 indicating a very good performance of reflectance measurements to correctly differentiate among the groups. We evaluated DRS spectra from 400 to 800 nm, it is likely that instead of using a full spectrum some specific spectral features might be sufficient to provide information about fat contents in the hepatic parenchyma. For example, key wavelengths like 560 nm might offer valuable information since it coincides with Hb and fat absorbance, and our data show an increase in reflectance associated to steatosis. Other features could also contribute, including the AUC not only for the full spectra but also for intervals of characteristic regions such as 500 to 570, 570 to 635, and 635 to 800 nm, all of which exhibited a significant difference between the means of the groups at the 0.05 level. Previously, we discussed the slopes in the region from 500 to 550 nm; however, the uncertainties of the slope values were too high to be used as a classification parameter for steatosis, at least in our model. Some other models of NAFLD could provide valuable information regarding the use of DRS in staging this disease and even differentiating among micro- and macrovesicular steatosis, including the genetically modified mouse (ob/ob) or rat (fa/fa); or the metabolically affected by diet induced obesity, etc. The MCD model of NAFLD used here shows micro-and macrovesicular steatosis in the early stages (MCD2w) but as steatosis advances, only macrovesicular steatosis is observed (MCD8w). Our results suggest that the size of the fat droplets could contribute to the increase of the observed reflectance.

Some limitations in our study include the use of experimental samples; however, with the evidence shown here we would be able to apply this tool in assessing samples of clinical impact such as liver grafts and surgical biopsies. Also we studied only simple steatosis; inflammation and ballooning in NASH are still to be analyzed by DRS.

Another limitation was the invasiveness of our approach; noninvasive alternatives have been developed in order to assess liver steatosis. For example, application of controlled attenuated parameter (CAP) in transient elastography is able to differentiate among degrees of steatosis.[33] CAP measures the ultrasonic attenuation in the liver based on the fact that fat affects ultrasound propagation.[34] Although CAP is highly promising for assessing steatosis, it is unable to differentiate micro- from macrovesicular steatosis and simple steatosis from NASH; therefore, more tools need to be developed.

Further research should be conducted to successfully apply DRS in the diagnosis of liver steatosis to discern simple steatosis from NASH, or NASH from developing fibrosis. Although we

measured *ex vivo* livers through the Glisson's capsule, we do not consider it as a limitation since it has been shown that surface measurements through the capsule are representative of the liver parenchyma.[16]

5 Conclusions

We have shown in an animal model that DRS with vis-NIR light is a highly sensitive and specific tool for the evaluation of steatosis. The systemic fat accumulation in the liver raises the diffuse reflectance. That is particularly evident in the wavelength range of 550 to 800 nm. The area under the diffuse reflectance curve reliably differentiates MCD2w and MCD8w groups where fat histopathological evaluation method fails. The multivariate statistical analysis shows that it is possible to use DRS measurements to accurately categorize the different stages of steatosis. The high correlation between the reflectance measurements and the fat histopathological evaluation indicates that the DRS could fill the existing void in evaluation of the type and degree of steatosis.

References

1. A. Lonardo et al., "Nonalcoholic fatty liver disease: a precursor of the metabolic syndrome," *Dig. Liver Dis.* **47**(3), 181–190 (2015).
2. L. Calzadilla Bertot and L. A. Adams, "The natural course of non-alcoholic fatty liver disease," *Int. J. Mol. Sci.* **17**(5), 774 (2016).
3. P. Bedossa, "Current histological classification of NAFLD: strength and limitations," *Hepatol. Int.* **7**(Suppl 2), 765–770 (2013).
4. J. M. Monteiro et al., "Nonalcoholic fatty liver disease: different classifications concordance and relationship between degrees of morphological features and spectrum of the disease," *Anal. Cell Pathol. (Amst)* **2014**, 526979 (2014).
5. E. K. Spengler and R. Loomba, "Recommendations for diagnosis, referral for liver biopsy, and treatment of nonalcoholic fatty liver disease and nonalcoholic steatohepatitis," *Mayo Clin. Proc.* **90**(9), 1233–1246 (2015).
6. O. Pournik et al., "Inter-observer and intra-observer agreement in pathological evaluation of non-alcoholic fatty liver disease suspected liver biopsies," *Hepat. Mon.* **14**(1), e15167 (2014).
7. M. Munteanu et al., "Diagnostic performance of FibroTest, SteatoTest and ActiTest in patients with NAFLD using the SAF score as histological reference," *Aliment. Pharmacol. Ther.* **44**(8), 877–889 (2016).
8. Y. Sumida, A. Nakajima, and Y. Itoh, "Limitations of liver biopsy and non-invasive diagnostic tests for the diagnosis of nonalcoholic fatty liver disease/nonalcoholic steatohepatitis," *World J. Gastroenterol.* **20**(2), 475–485 (2014).
9. D. A. Fabila-Bustos et al., "Diffuse reflectance spectroscopy as a possible tool to complement liver biopsy for grading hepatic fibrosis in paraffin-preserved human liver specimens," *Appl. Spectrosc.* **68**(12), 1357–1364 (2014).
10. K. Vishwanath et al., "Portable, fiber-based, diffuse reflection spectroscopy (DRS) systems for estimating tissue optical properties," *Appl. Spectrosc.* **65**(2), 206–215 (2011).
11. R. S. DaCosta, H. Andersson, and B. C. Wilson, "Molecular fluorescence excitation-emission matrices relevant to tissue spectroscopy," *Photochem. Photobiol.* **78**(4), 384–392 (2003).
12. D. Fabila et al., "*In vivo* assessment of liver fibrosis using diffuse reflectance and fluorescence spectroscopy: a proof of concept," *Photodiagn. Photodyn. Ther.* **9**(4), 376–382 (2012).
13. V. Tuchin, *Tissue Optics, Light Scattering Methods and Instruments for Medical Diagnostics*, 3rd ed., SPIE Press, Bellingham, Washington (2015).
14. N. Reistad et al., "Diffuse reflectance spectroscopy of liver tissue," *Proc SPIE* **9531**, 95314E (2015).
15. D. J. Evers et al., "Diffuse reflectance spectroscopy: toward real-time quantification of steatosis in liver," *Transpl. Int.* **28**(4), 465–474 (2015).
16. J. H. Nilsson et al., "Diffuse reflectance spectroscopy for surface measurement of liver pathology," *Eur. Surg. Res.* **58**(1–2), 40–50 (2017).
17. A. C. Westerkamp et al., "Diffuse reflectance spectroscopy accurately quantifies various degrees of liver steatosis in murine models of fatty liver disease," *J. Transl. Med.* **13**, 309 (2015).
18. D. E. Kleiner et al., "Design and validation of a histological scoring system for nonalcoholic fatty liver disease," *Hepatology* **41**(6), 1313–1321 (2005).
19. J. M. Miller, J. C. Miller, and R. D. Miller, *Statistics and Chemometrics for Analytical Chemistry*, 6th ed., Pearson Education Limited, Harlow, England (2010).
20. T. F. Hastie, R. Tibshirani, and J. H. Friedman, *The Elements of Statistical Learning: Data Mining, Inference, and Prediction*, 2nd ed., Springer, New York (2009).
21. S. J. Dixon and R. G. Brereton, "Comparison of performance of five common classifiers represented as boundary methods: euclidean distance to centroids, linear discriminant analysis, quadratic discriminant analysis, learning vector quantization and support vector machines, as dependent on data structure," *Chemometr. Intell. Lab. Syst.* **95**(1), 1–17 (2009).
22. R. Kumar and A. Indrayan, "Receiver operating characteristic (ROC) curve for medical researchers," *Indian Pediatr.* **48**(4), 277–287 (2011).
23. M. D. Ruopp et al., "Youden index and optimal cut-point estimated from observations affected by a lower limit of detection," *Biometrical J.* **50**(3), 419–430 (2008).
24. R. Nachabe et al., "Effect of bile absorption coefficients on the estimation of liver tissue optical properties and related implications in discriminating healthy and tumorous samples," *Biomed. Opt. Express* **2**(3), 600–614 (2011).
25. A. N. Bashkatov et al., "Optical properties of human skin, subcutaneous and mucous tissues in the wavelength range from 400 to 2000 nm," *J. Phys. D Appl. Phys.* **38**(15), 2543–2555 (2005).
26. N. Bosschaart et al., "A literature review and novel theoretical approach on the optical properties of whole blood," *Lasers Med. Sci.* **29**(2), 453–479 (2014).
27. R. Graaff et al., "Reduced light-scattering properties for mixtures of spherical-particles—a simple approximation derived from Mie calculations," *Appl. Opt.* **31**(10), 1370–1376 (1992).
28. P. Parsa, S. L. Jacques, and N. S. Nishioka, "Optical-properties of rat-liver between 350 and 2200 nm," *Appl. Opt.* **28**(12), 2325–2330 (1989).
29. S. L. Jacques, "Optical properties of biological tissues: a review," *Phys. Med. Biol.* **58**(11), R37–R61 (2013).
30. U. S. Dinish et al., "Diffuse optical spectroscopy and imaging to detect and quantify adipose tissue browning," *Sci. Rep.* **7**, 41357 (2017).
31. D. A. Fabila-Bustos et al., "Fluorescence spectroscopy as a tool for the assessment of liver samples with several stages of fibrosis," *Photomed. Laser Surg.* **36**(3), 151–161 (2018).
32. T. M. Bydlon et al., "Chromophore based analyses of steady-state diffuse reflectance spectroscopy: current status and perspectives for clinical adoption," *J. Biophotonics* **8**(1–2), 9–24 (2015).
33. R. P. Myers et al., "Controlled attenuation parameter (CAP): a noninvasive method for the detection of hepatic steatosis based on transient elastography," *Liver Int.* **32**(6), 902–910 (2012).
34. M. Sasso et al., "Controlled attenuation parameter (CAP): a novel VCTE guided ultrasonic attenuation measurement for the evaluation of hepatic steatosis: preliminary study and validation in a cohort of patients with chronic liver disease from various causes," *Ultrasound Med. Biol.* **36**(11), 1825–1835 (2010).

Eduardo J. Arista Romeu received his engineering degree in electronics and his master's degree from the Kiev Polytechnic Institute (KPI) in 1983 and his second master's degree from the ISPJAE (Havana Polytechnic Institute) in 2012. He is a PhD student in electronics and communications at the National Polytechnic Institute in Mexico. His current research interests focus on optical spectroscopy and data analysis. Prior to this, he worked at the CEADEN in Havana, where he specialized in the development of nuclear electronics for health applications and laser instrumentation.

Galileo Escobedo received his PhD in biomedicine in 2008. Currently, he is a PI at the General Hospital of Mexico. His main research interests are diagnosis of liver fibrosis and steatosis, as well systemic inflammation as a key contributor of chronic liver disease.

Adriana Campos-Espinosa received her bachelor's degree in biological-pharmaceutical chemistry from the Universidad Nacional Autónoma de México in Mexico City in 2018. She has been collaborating in basic research projects related to physiopathology of nonalcoholic fatty liver disease (NAFLD). She was awarded first place in a research competition for her project of lipid quantification in a murine model of NAFLD.

Ivette Irais Romero-Bello received her bachelor's degree in chemistry from the School of Chemistry, National University of Mexico (UNAM) in 2015. Later, she received her master's degree in clinical biochemistry from UNAM in 2018.

Javier Moreno-González received his bachelor's degree in chemistry from the School of Chemistry, National University of Mexico (UNAM) in 2016.

Diego A. Fabila-Bustos received his MSc and PhD degrees in electrical engineering from the Instituto Politecnico Nacional (IPN) in Mexico City in 2010 and 2014, respectively. He is a full professor at Unidad Profesional Interdisciplinaria de Ingeniería Campus Hidalgo (UPIIH) at the IPN. His current research interest are biomedical optics, photonics, and electronics instrumentation.

Alma Valor Reed received her doctorate degree in advanced technology from the Instituto Politécnico Nacional, and her master's degree in physics and mathematics from Moscow State University. She is a full professor at the Higher School of Mechanical and Electrical Engineering, Instituto Politécnico Nacional (National Polytechnic Institute), Zacatenco, México. Her current research interests include mathematical statistics, light–matter interaction modeling, and biophotonics.

Suren Stolik Isakina received his master's degree in physics and mathematics from Moscow State University and his PhD in advanced technology from the Instituto Politécnico Nacional. He is a full professor at Escuela Superior de Ingeniería Mecánica y Eléctrica, Instituto Politécnico Nacional, Zacatenco, México. His current research interests include biomedical optics, light–matter interaction, biophotonics, and photonics instrumentation.

José Manuel de la Rosa Vázquez received his BS and MSc degrees in electrical engineering from the National Polytechnic Institute (IPN) in México City in 1979 and 1981, respectively, and his Dr. Ing. degree from the Optics Institute of the Technical University Berlin in 1987. Since 1988, he has been a full professor at the IPN and his current research interests are in the biophotonics and bioelectronics areas.

Carolina Guzmán received her bachelor's degree in chemistry in 2001 and her PhD in biomedicine in 2007 from the National University of Mexico (UNAM). She was awarded the L'Oréal-UNESCO-AMC-For Women in Science prize in 2012. She is a researcher at the Hospital General de México, and a professor at the School of Chemistry, UNAM. Her research interests are diagnosis and physiopathology of chronic liver diseases.